Snow Cover Mapping Using Satellite Remote Sensing Data

[1]Hosni Ghedira, [2]Juan-Carlos Arevalo, [3]Tarendra Lakhankar, [3]Reza Khanbilvardi and [3]Reginald Blake
[1]Masdar Institute of Science and Technology, AbuDhabi, UAE
[2]Baker Engineering, Inc. 245 Fifth Avenue, New York, USA
[3]NOAA-CREST Center, City University of New York, USA
tlakhankar@ccny.cuny.edu

***Abstract* -This paper discusses neural network based approach to generate the spatial distribution of snow accumulation using multi-channel Special Sensor Microwave/Imager (SSM/I) data. Five SSM/I channels (19H, 19V, 22V, 37V, and 85V) were used to remotely sense snow accumulation during 2001/2002 winter season. Ground snow depth measurements were acquired from the National Climatic Data Center (NCDC) through the Cooperative Observer Network for snow monitoring in the United States. The snow depths were compiled and gridded into 25 km x 25 km grid to match the final SSM/I spatial resolution. Neural network based approach was tested and compared with the filtering algorithm developed by Grody and Basist[1] in the Northern Midwest region of the United States. The results indicate that the neural-network-based approach has a great potential in identifying snow pixels from SSM/I data by providing a significant improvement in snow mapping accuracy over the filtering algorithm.**

Keywords:Snow mapping; SSM/I; Passive Microwave; Neural Network.

I. INTRODUCTION

Having accurate estimations of snow cover characteristics during the snowmelt season is indispensable for efficient hydrological modeling and snowmelt runoff forecasting [2]. Direct measurements of snow depth at a single station are generally not very useful in making estimates of accumulation over large areas. Additionally, the traditional field sampling methods and the ground-based data collection are often very sparse, time consuming, and expensive compared to the coverage provided by remote sensing techniques. Moreover, direct measurements of snow depth at a single station are generally not very useful in making estimates of distribution over large areas since the measured depth may be highly unrepresentative of the study areas even under the same snowfall conditions. At present, most hydrological models that require snowpack information are using maps obtained by gridding standard point gauge measurements or data derived from physically based models [2-4].

The estimation of snow depth and snow water equivalent from passive microwave measurements requires a deep understanding of surface and volume emissivity of snowpack and its underlying ground. The measured brightness temperature of the snow-covered surface is a function of both ground and snow cover properties, includes: surface roughness, surface temperature, vegetation cover, snow cover density, snow water equivalent, and snow grain size distribution [3].

Many empirical models have been developed to estimate snow depth from spaceborne passive microwave sensors; most of these models make the simple assumption that the snow depth and brightness temperature differences, generally between channels 19 and 37 GHz, are linearly related. The Meteorological Service of Canada (MSC) model, for example, currently uses them to produce real-time SWE maps for the Canadian Prairies [5, 6]. In forest environments, SWE retrieval becomes more complicated due to the attenuation of the ground microwave signal propagating through the canopy as well as the vegetation contribution to the brightness temperature [7, 8].

Neural network has been successfully applied to a wide range of non-linear problems in several disciplines. Multi-layer perceptron trained by the backpropagation algorithm has also been successfully applied to image classification, and it has shown great potential in the classification of different types of remotely sensed data. A useful review of the application of neural networks in remote sensing can be found in [9, 10]. The rapid increase in neural network applications in remote sensing is mainly due to their ability to perform more accurately than other classification techniques especially when the intent is to classify features with overlapped spectral signatures that cannot be easily associated with defined statistical functions. Generally, a neural network is capable of storing a complex functional relationship between its inputs (pixel values) to the outputs and it is proficient in approximating any function with a finite number of discontinuities.

The major advantage of the neural network over traditional classifiers is its easy adaptation to different types of data and input configurations (decimal or binary). Moreover, neural networks can easily incorporate ancillary data sources which would be difficult to integrate with conventional techniques. Traditional parametric classification methods such as: Maximum Likelihood Classifier makes unreasonable assumptions about the statistical properties of the data. These assumptions are not always satisfied especially when heterogeneous natural land covers are being considered. Such assumptions are avoided by the neural network. A neural network uses its complex configuration to find the best nonlinear function between the input and the output data without the constraint of linearity or pre-specified non-linearity, which is required in regression analysis. Unlike most statistical classification methods, neural networks have

the capacity to weigh differently and automatically each data source according to its contribution to ground cover identification [9].

In this study, neural network was used to retrieve the spatial distribution of snow accumulation from multi-channel SSM/I data. The application of neural network in snow mapping was compared to filtering algorithm developed by Grody and Basist[1]. Neural network training approaches based on snow depth were investigated by varying the selection criteria training pixels to improve snow cover classification accuracy.

II. STUDY AREA AND DATA ACQUISITION

The study area is located in the Northern Midwest of the United States within 110˚63’W - 102˚04’W and 48˚71’N - 40˚73’N. The study area selection was based on the existence of a large number of meteorological stations with high snow accumulation. The passive microwave data from the Special Sensor Microwave/Imager (SSM/I) Level 3 EASE-Grid Brightness Temperatures was used in both ascending and descending orbits. These images provided measurements of the brightness temperature in seven channels with different frequencies and polarizations. In this study, the same five SSM/I channels (19H, 19V, 22V, 37V, and 85V) used by Grody and Basist[1] for the filtering algorithm have been used to train and validate the neural network algorithm.

Three post storm days with deep snow cover have been selected during the 2001/2002 winter season (01/23, 01/24, and 01/25). A total of 185 ground stations within the study area have been used for this experiment. The snow depth collected from these ground-stations was linearly interpolated to a regular grid over the study area to serve as truth data. The study area contains 34 x 30 pixels with spatial resolution of 25 km.

III. METHODOLOGY

A. Decision Tree based Filtering Algorithm

The microwave emissions between low and high-frequency channels varied based on snow cover and non snow pixels. A filtering algorithm based on microwave scattering theory for snow cover identification developed by Grody and Basist[1] was used in this study. This algorithm uses physical relationship between the measured brightness temperature at frequencies and scattering response to snow cover or precipitation.

The algorithm consists of a decision tree which establishes sensitive thresholds to filter out precipitation, cold desert and frozen surfaces (Figure 1). This filtering algorithm separates the scattering signature of snow from the scattering signatures of precipitation, cold deserts, and frozen ground. This filtering algorithm uses the antenna temperature retrieved from five SSM/I channels (19V, 19H, 22V, 37V, and 85H). More details about this technique can be found in [1].

B. Neural Network

A Neural network is a highly interconnected system of simple processing elements (called nodes) that is designed to mimic the highly parallel human biological neurons. These nodes are usually organized into a sequence of layers with random connections between successful layers. The strength of these connections is given through the connecting weights of the network. Each node calculates a summation of weighted inputs and then outputs its transfer function value to other nodes. A multi-layer neural network consists of a number of interconnected nodes with each node operating as a simple processing element. The processing nodes are arranged in layers. Each node is interconnected with all nodes in the preceding and following layers. There are no interconnections within nodes of the same layer. The number of layers and the number of nodes by layer represent the network architecture. The input layer serves as an entry for the vector of data presented to the network (SSM/I channels). The output layer serves to produce the neural network decision (snow or non-snow) for the data presented at the input layer. All layers between the input and output layers are referred to hidden layers. The input (Ij) to a node (j) is the weighted sum of the outputs (Oi) from the nodes of the preceding layer (i). This sum is then passed through an activation function (f) to produce the node’s output (Oj) within the range of the activation function. The activation function is usually a sigmoid or hyperbolic tangent, which are non-linear functions that have an asymptotic behavior. The best neural network architecture can only be determined experimentally for each particular problem. The number of hidden nodes should be large enough to ensure a sufficient number of degrees of freedom for the network function and simultaneously small enough to keep sufficient the generalization ability to the network [11].

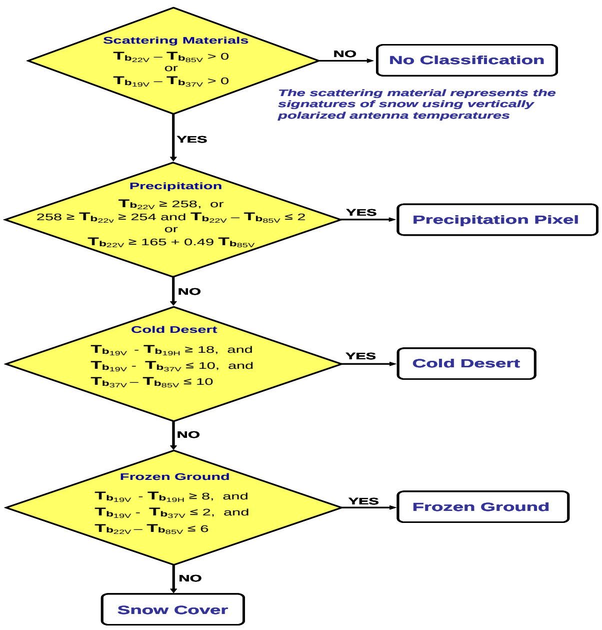

Figure 1. Decision Tree for global snow cover identification adapted from [1].

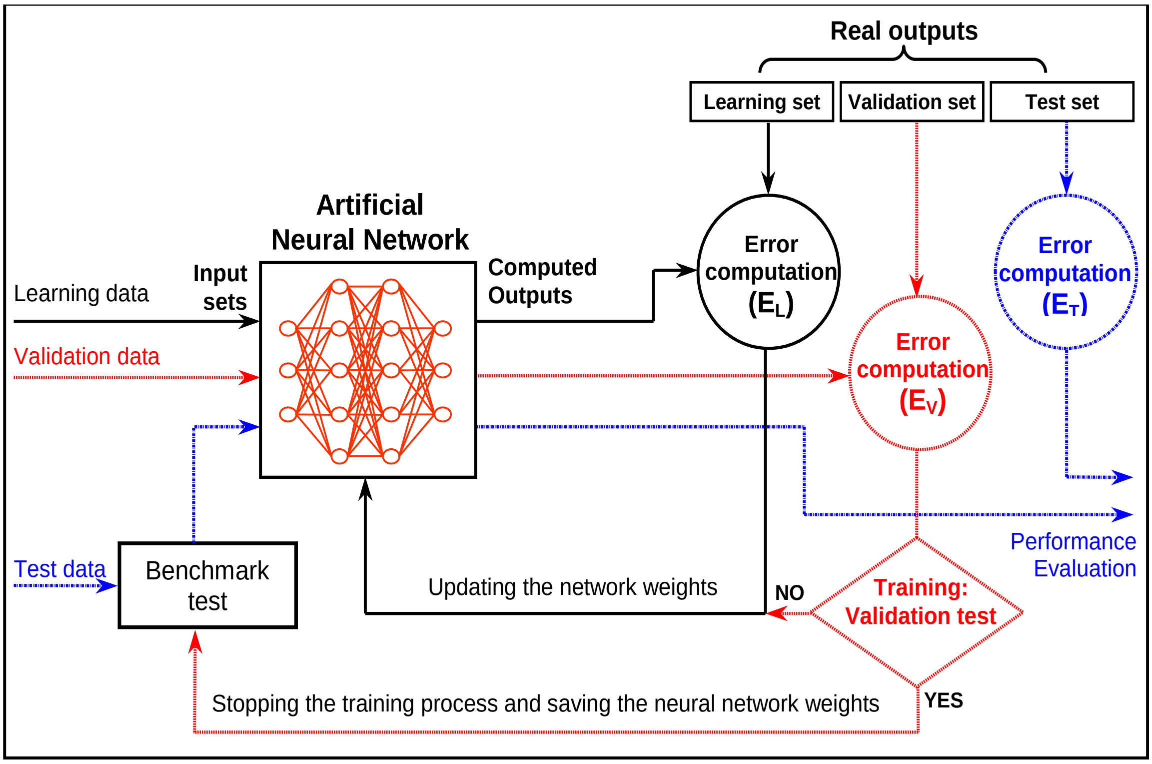

Figure 2.Contribution of each training group to the global training process

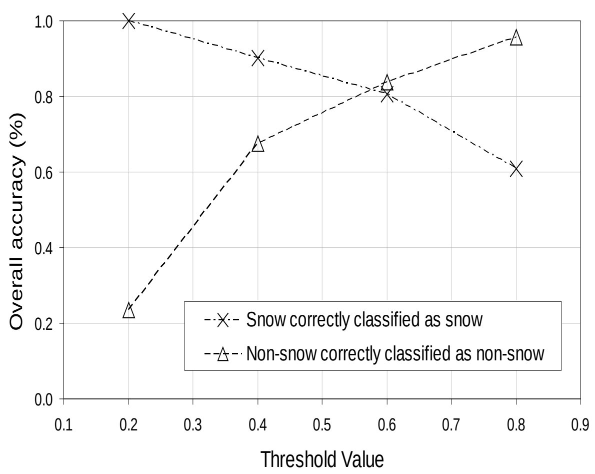

Figure 3. Effect of threshold values on snow and non-snow classification accuracy

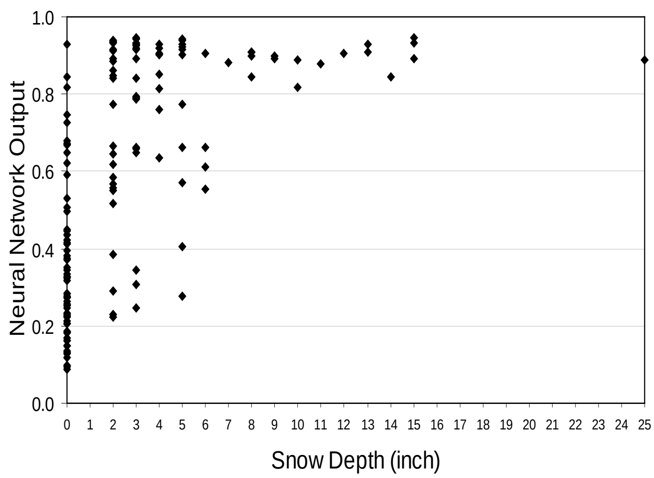

Figure 4. Neural network classification accuracy compared as a function of snow depth

C. Neural Network Training Process

The training stage consists in adjusting the connection weights (randomly initialized) in order to decrease the difference between the network output and the desired output. The training data were presented to the input layer and propagated through the hidden layer to the output layer. The differences between the neural network outputs and the desired outputs were computed and feed-backward to adjust the network connections. One of the major concerns in the neural network training process is overtraining. When overtraining occurs, the neural network's generalization ability will be compromised and the classification space becomes narrowly defined around the training pixels [12]. To avoid

overtraining of the neural network, the available training data were divided into three subsets. The first subset was the training set; this set was used for computing and updating the network weights. The second subset, the validation set, was used to avoid the overtraining of the neural network by monitoring the validation error during the training process. The third subset, the test set, was not used during the training process. The test set was only used to benchmark the neural network and to compare different models. The methodology used in the training process is illustrated in the Figure 2. Normally, as it is the case for the training set error, the error computed on the validation set decreases during the initial phase of training. However, when the network begins to overfit the training data, the error on the validation set will begin to increase slowly for the next iterations. At that time, the training process will be stopped, and the neural network weights corresponding to the minimum validation error will be maintained for the testing neural network performance.

In this study, for each vector of five brightness temperatures presented in the input layer, a value equal to one will be assigned in the output layer if the presented vector corresponds to a snow pixel. Otherwise, a value equal to zero will be assigned. However, due to the asymptotic behavior of the activation function, a continuous range from zero to one was produced by the output neuron during the simulation process. This variability can be explained by the fact that the neural network could not be trained to produce a zero error on the training data, additionally, the data being classified could also be more diversified than the data used in the training. Based on several runs of neural networks, we have observed no apparent advantage for multi-hidden-layer networks over single-hidden-layer networks for our data. Thus, we used a single hidden layer network structure (5:10:1) to predict the snow cover in this study. Network architecture 5:10:1 was used in the next processing steps.

D. Neural Network Testing Approach

To transform the continuous output format into a categorical format, a threshold value between 0 and 1 has been introduced to decide if the pixel will be labeled as snow or non-snow. The optimal threshold value cannot be identified with certainty without measuring its effect on the classification accuracy. In this project, the threshold value has been varied from 0.2 to 0.8. The effect of the decision threshold on classification accuracy of each class is illustrated in Figure 3. This figure shows that the threshold value affects the overall classification significantly. Specifically, the increase of the threshold value results in a simultaneous decrease of the percentage of correctly classified snow pixels and an increase in the percentage of correctly classified non-snow pixels. For this specific training, a threshold of 0.6 has been retained providing an overall accuracy of 75, 76 and 78% respectively on the test set for 3 consecutive testing days.

The scatter plot (Figure 4) shows more precisely the crucial role of the threshold selection in providing an accurate classification and how the neural network performs better for the pixels with high snow accumulations. Indeed, for all the pixels with snow depth higher than 6 inches, the neural network accuracy was higher than 0.8. Thus, if we select a threshold equal to 0.8, all the pixels with snow accumulation higher than 6 inches will be correctly classified and only 3 non-snow pixels will be misclassified.

E. Data Feeding to Neural Network Training

The proper selection of training data is a crucial step to achieve best results. To ensure an accurate selection of training pixels, four approaches were tested by varying the selection criteria of snow pixels.

1. First approach: all the pixels with one inch or more of snow accumulation were considered as snow pixels.
2. Second approach: only the pixels with two inches or more of snow depth were considered as snow pixels. This approach reduces the risk of overestimating the ground snow depth during the interpolation (or gridding) of the snow gauge measurements. In this approach, the neural network was trained to classify the one-inch snow pixels as no-snow pixels.
3. Third approach: all the pixels with one inch of snow depth have been removed from the training process to reduce the risk of mislabeling them as snow or non-snow pixels.
4. Fourth approach: only the pixels with ground stations inside their boundaries were used for the training. For these pixels, only those with accumulation larger than one-inch were considered as snow pixels.

By comparing the four approaches, we find that the third and the fourth approaches give the better performances by reducing the misclassification of non-snow pixels by about 10% and increasing the accuracy of correctly classified snow pixels by about 5%. The third and fourth approach give better performance could be attributed to reduction in SSM/I sensitivity shallow dry snow and melting snow [1].

IV. Results and Discussion

A comparison between the neural network technique and the filtering algorithm method was conducted. The third set (200 pixels), which was not included in training process was used to test the performance of neural network. The performance was tested in terms of calculating confusion matrices. Confusion matrices were calculated for neural network and filtering algorithm for the three selected days. The confusion matrices presented in Table 1 show that the neural network (trained using approach 3) provides a significant improvement in snow mapping accuracy over the filtering algorithm. However, the filtering algorithm slightly outperforms the neural network by 2 percent on one day (Jan 25). In terms of categorical assessment, higher Kappa coefficient [13] values were observed for Jan 23 and Jan 24 using neural network method compared to filtering algorithm.

The snow maps shown in Figure 5 represent the gridded gauge measurements and the output of each technique with the same inputs for the three selected days. The decision tree maps represent the output of the filtering algorithm, and the neural network maps represent the simulation results of the selected neural network (threshold = 0.6 and training by approach 3). In order to have a consistent comparison between the two

techniques, the same five channels used in the filtering algorithm were used to train and simulate the neural network technique.

Date	Row	Decision Tree S	Decision Tree NS	ANN S	ANN NS
Jan 23	S	**0.86**	0.69	**0.57**	0.20
	NS	0.14	**0.31**	0.43	**0.80**
		Accuracy =43 Kappa =9.2		Accuracy =75 Kappa =33.3	
Jan 24	S	**0.70**	0.48	**0.38**	0.16
	NS	0.30	**0.52**	0.62	**0.84**
		Accuracy =56 Kappa =1.3		Accuracy =76 Kappa =21.6	
Jan 25	S	**0.56**	0.15	**0.57**	0.17
	NS	0.44	**0.85**	0.43	**0.83**
		Accuracy =80 Kappa =38.1		Accuracy =78 Kappa =35.9	

Table 1. Filtering algorithm (Decision Tree) and Neural Network (ANN) performance assessment using confusion matrix

(S = Snow and NS = No Snow)

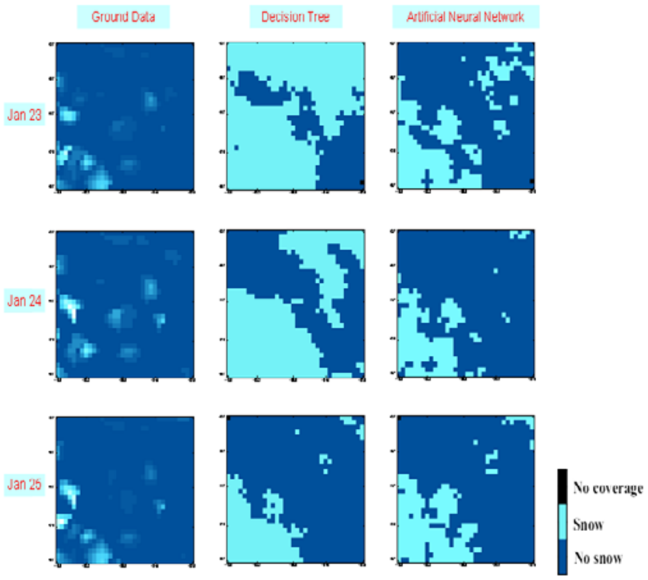

Figure 5. Comparison of the ground truth data with the decision tree and neural network outputs for Northern Midwest of the United States within 110°63'W - 102°04'W and 48°71'N - 40°73'N.

Overall, the performance of the neural network in identifying snow-covered pixels was around 77 percent. This accuracy is approximately 15 percent higher (on average for 3 days of data) than the one obtained with the filtering algorithm when tested on an independent set of data (not used in the training and validation of neural network). Furthermore, for shallow snow cover, misleading results were obtained when the one-inch snow pixels were considered as snow pixels. It was found that for deeper snow depth (>2 inches), the snow and non-snow pixels were more likely to be correctly classified when a threshold of 0.4 or 0.6 was used. The same results showed that the mapping accuracy is positively correlated with the snow depth. The low accuracy of shallow snow covers can be explained by the mislabeling of whole pixels (one pixel covers approximately 625 km^2) as snow-covered by using only one or two point observations that record less than 2 inches of snow. An attempt to overcome this source of error was made by removing low-accumulating pixels (less than one inch) from the neural training (approach 3). This approach showed a slight improvement of the overall accuracy (less than 5 percent), but most of the pixels with shallow snow cover were still misclassified during the spatial simulation of the trained neural network over large areas. Such misclassification could be reduced by using other sources of truth data (aerial or satellite-based maps acquired under cloud free conditions) instead of using interpolated point measurements.

V. CONCLUSION

This study explores the ability of neural networks to improve the mapping of snow cover using SSM/I data. The results indicate that neural networks can be considered as an alternative to retrieve snow cover information from passive microwave satellite measurements. Four neural network approaches based on snow depth were tested by varying the selection criteria training pixels to improve snow cover classification accuracy. This was attempted to overcome the source of error by removing low-accumulating (thin snow) pixels from the neural network training. The use of non-parametric tools such as neural network facilitates the representation of the true processes by simpler parametric relations. The neural network does not provide relationships between variables and considered as a black box. However, neural network help to identify significant variables in the system for which physical relationship can be developed. This study focused on the United States because a large quantity of in-situ gauge data was readily available for this area.

ACKNOWLEDGEMENTS

This study was supported and monitored by National Oceanic and Atmospheric Administration (NOAA) under Grant NA06OAR4810162. The views, opinions, and findings contained in this report are those of the author(s) and should not be construed as an official National Oceanic and Atmospheric Administration or U.S. Government position, policy, or decision.

REFERENCES

[1] N. C. Grody and A. N. Basist, "Global identification of snowcover using SSM/I measurements," IEEE Transactions on Geoscience and Remote Sensing, vol. 34, pp. 237-249, 1996.

[2] R. C. Bales and R. F. Harrington, "Recent progress in snow hydrology," Reviews of Geophysics, vol. 33(S1), pp. 1011–1020, 1995.

[3] A. Rango, "Progress in Snow Hydrology Remote-Sensing Research," IEEE Transactions on Geoscience and Remote Sensing, vol. GE-24, pp. 47-53, 1986.

[4] T. Lakhankar, et al., "Analysis of Large Scale Spatial Variability of Soil Moisture Using a Geostatistical Method," Sensors, vol. 10, pp. 913-932, 2010.

[5] B. E. Goodison, "Determination Of Areal Snow Water Equivalent On The Canadian Prairies Using Passive Microwave Satellite Data," in Geoscience and Remote Sensing Symposium, 1989. IGARSS'89. 12th Canadian Symposium on Remote Sensing., 1989 International, 1989, pp. 1243-1246.

[6] A. E. Walker and B. E. Goodison, "Discrimination of a wet snowcover using passive microwave satellite data," Annals of Glaciology, vol. 17, pp. 307-311, 1993.

[7] K. Goïta, et al., "Estimation of snow water equivalent in the boreal forest using passive microwave data," Proc. International Symposium Geomatics in the Era of Radarsat, May 26-30 1997, Ottawa, Canada, 1997.

[8] K. Goita, et al., "Algorithm development for the estimation of snow water equivalent in the boreal forest using passive microwave data," International Journal of Remote Sensing, vol. 24, pp. 1097 - 1102, 2003.

[9] J. Benediktsson, et al., "Neural network approaches versus statistical methods in classification of multi-source remote sensing data," IEEE Transactions on Geoscience and Remote Sensing, vol. 28, pp. 540-552, 1990.

[10] J. D. Paola and R. A. Schowengerdt, "A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery," International Journal of Remote Sensing, vol. 16, pp. 3033 - 3058, 1995.

[11] D. E. Rumelhart, et al., Learning internal representation by error propagation, Parallel distributed processing: Exploration in the microstructure of cognition: M.I.T., Cambridge Press (MA), 1986.

[12] M. Augustejin, et al., "Performance evaluation of texture measures for ground cover identification in satellite images by means of a neural network classifier," IEEE Transactions on Geoscience and Remote Sensing, vol. 33, pp. 616-626, 1995.

[13] S. V. Stehman, "Estimating the Kappa Coefficient and its variance under stratified random sampling," Photogrammetric Engineering & Remote Sensing, vol. 62, pp. 401-407, 1996.